

Human actions recognition on multimedia hardware using angle-based and coordinate-based features and multivariate continuous hidden Markov model classifier

Tomasz Hachaj¹ · Marek R. Ogiela²

Received: 14 December 2014 / Revised: 5 July 2015 / Accepted: 28 August 2015 /

Published online: 5 September 2015

© The Author(s) 2015. This article is published with open access at Springerlink.com

Abstract In this paper we have proposed human actions recognition schema that uses angle-based and state-of-the-art coordinates-based features and multivariate continuous hidden Markov model (HMM) classifier with Gaussian distribution. The main novelty of this paper besides presenting this approach is its evaluation on large (containing 770 actions) motion-capture dataset of various gym exercises. We have evaluated HMM with various number of hidden states and Gaussian mixture model classifier. We have performed PCA analysis of both features sets to justify our choices. We have also test 24 subsets of proposed angle-based features dataset in order to estimate which body joints are vital for correct actions recognition. The highest recognition rate in k-fold cross validation was 97 ± 14 % and was obtained for 4-states HMM with 10 angle-based features. The common problem with markerless motion capture vision - based systems is that if some parts of body surface are covered by another part it is impossible to perform accurate features measurements. Knowing this we have evaluated 24 subsets of proposed features dataset in order to estimate which body joints are vital for correct actions recognition.

Keywords Human actions recognition · Hidden Markov models · Features selection · Classification · Gym exercises · Kinect

✉ Tomasz Hachaj
tomekhachaj@o2.pl

Marek R. Ogiela
mogiela@agh.edu.pl

¹ Institute of Computer Science and Computer Methods, Pedagogical University of Krakow, 2 Podchorążych Ave, 30-084 Krakow, Poland

² Cryptography and Cognitive Informatics Research Group, AGH University of Science and Technology, Mickiewicza Ave, 30-059 Krakow, Poland

1 Introduction

The classification of human body movements is very interesting and still open problem. There are numbers of common and specialized activities a human do every day. There are many reasons why analysis of those actions is in the scope of interest of science and industry. For example properly classified hand gestures or full body movements can be used as a user interface for digital systems. Human actions during training can be interpreted by computer coaching systems. Pattern recognition methods can also detect potentially hazardous situation that might lead to violence. In literature there are many approach that can be used to hand gestures or body action recognition (see state of the art in [11, 12]). However hidden Markov models (HMM) are among most popular are reliable methods for those tasks.

1.1 State of the art on application of HMM to gestures and actions recognition

Discrete HMM together with other pattern recognition techniques like fuzzy logic or neural networks (NN) [9, 10] has many application in classification of for example handwritten text [22], speech [25, 36] or gestures / actions [19, 29, 34]. The HMM classifier is often accompanied with another pattern recognition method that is used for discretization of continues signal. Those can be for example NN [20, 25] or Support Vectors Machine [36].

Many up-to date gestures and action recognition methods utilizes multi-dimensional continuous HMM, the commonly used distribution to describe the observation densities is the Gaussian one. The characteristics and dimensionality of input signal differs depending of data acquisition method. What is more very often a classifier requires pre-processing of input data.

For example in [19] authors present a glove-based hand gesture recognition. They decompose the input signal to a set of so called hand strokes. A stroke can be a simple hand gesture but usually two or three strokes constitute a more complex gesture. Once the discrete stroke HMMs have been trained for the defined strokes, hand gesture recognition can be accomplished by identifying a composition of such strokes.

In paper [34] a human activity recognition method is proposed, which utilizes Independent Component Analysis (ICA) for activity shape information extraction from image sequences and hidden Markov model for recognition. Various human activities are represented by shape feature vectors from the sequence of activity shape images via ICA. Based on these features, each HMM is trained and activity recognition is achieved by the trained HMMs of different activities.

Paper [13] introduces a model-based hand gesture recognition system, which consists of three phases: feature extraction, training, and recognition. In the feature extraction phase, a hybrid technique combines the spatial (edge) and the temporal (motion) information of each frame to extract the feature images. Then, in the training phase, authors use the principal component analysis (PCA) to characterize spatial shape variations and the hidden Markov models to describe the temporal shape variations. A modified Hausdorff distance measurement is also applied to measure the similarity between the feature images and the pre-stored PCA models. The similarity measures are referred to as the possible observations for each frame. Finally, in recognition phase, with the pre-trained PCA models and HMM, authors generate the observation patterns from the input sequences, and then apply the Viterbi algorithm to identify the gesture.

The Depth Map-Based Human Activity Recognition system [33] uses Silhouette vector extraction and dimensionality reduction with PCA to generate input for continuous hidden Markov models.

In [18] a hierarchical human activity recognition (HAR) system is proposed to recognize abnormal activities from the daily life activities of elderly people living alone. The system is structured to have two-levels of feature extraction and activity recognition. The first level consists of R-transform, kernel discriminant analysis (KDA), k-means algorithm and HMM to recognize the video activity. The second level consists of KDA, k-means algorithm and HMM, and is selectively applied to the recognized activities from the first level when it belongs to the specified group.

In [21] authors propose a forward spotting scheme that executes gesture segmentation and recognition simultaneously. The start and end points of gestures are determined by zero crossing from negative to positive (or from positive to negative) of a competitive differential observation probability that is defined by the difference of observation probability between the maximal gesture and the non-gesture. Authors also propose the sliding window and accumulative HMMs.

In the method proposed in [17] a gesture is represented as an ordered sequence of poses in a spatiotemporal space. To estimate a human pose in a frame, all frames extracted from an image sequence are classified. The features of a frame are represented by the position values of three upper body parts: head, left hand, and right hand. The further analysis is based on trajectories recognition.

Paper [14] introduces a multi-Principal-Distribution-Model (PDM) method and hidden Markov model for gesture recognition. To track the hand-shape, it uses the PDM model which is built by learning patterns of variability from a training set of correctly annotated images.

The method for hand gesture interpretation in Sign Language [8] uses the raw data that is captured using a color digital camera in the video format. Then a SwisTrack implementation of the color separation and tracking method is used to extract the 2D trajectories of head and hands of the signer. Extracted trajectories are converted to a Polar coordinate system from Cartesian coordinates at the first stage. Stationary head or hand components are not informative. Hence they are removed in the second stage. Then at stage three, it is possible to distinguish gestures with a single or both hands, assuming the head remains relatively stationary. In the following stage 4, the mean co-ordinate of the gesture is removed to reposition it at the origin. In stage 5, the gesture is rotated and scaled down to a neutral position using Singular Value Decomposition (SVD). Finally in stage 6, the gesture is re-sampled using Bezier curve techniques, in order to obtain trajectories that are jitter-free, clean and smooth, having standard sampling rate.

In paper [6] authors introduce a hand gesture recognition system to recognize continuous gesture before stationary background. The system consists of four modules: a real time hand tracking and extraction, feature extraction, hidden Markov model training, and gesture recognition. First, authors apply a real-time hand tracking and extraction algorithm to trace the moving hand and extract the hand region, then they use the Fourier descriptor (FD) to characterize spatial features and the motion analysis to characterize the temporal features. They combine the spatial and temporal features of the input image sequence as our feature vector. After having extracted the feature vectors, HMMs is applied to recognize the input gesture.

Paper [15] presents EasyGR (Easy Gesture Recognition), a tool based on machine learning algorithms that help to reduce the effort involved in gesture recognition. Authors use joint-

based user tracking. The input features are generated by normalization of body position to make it camera position invariant and HMM for recognition.

In research [35] a HMM is proposed for various types of hand gesture recognition. In the preprocessing stage, this approach consists of three different procedures for hand localization, hand tracking and gesture spotting. The hand location procedure detects hand candidate regions on the basis of skin-color and motion. The hand tracking algorithm finds the centroids of the moving hand regions, connects them, and produces a hand trajectory.

In paper [4] authors propose a method of modeling hand gestures based on the angles and angular change rates of the hand trajectories. Each hand motion trajectory is composed of a unique series of straight and curved segments. In hidden Markov model implementation, these trajectories are modeled as a connected series of states analogous to the series of phonemes in speech recognition.

In [24] authors use discrete HMM on derivatives of tracked body joints of shoulder, elbow and wrist. Those joints are used to calculate angle features of left and right arm. Afterward, the gesture feature vector is clustered using the k-means clustering algorithm [4], the k is set as 16.

Also [37] uses gesture grouping with a k-means algorithm before presenting the features to HMM. The input signal is gathered with three-dimensional finger worn accelerometer.

In the work [26] a background modeling algorithm using fuzzy logic is used to detect foreground objects in outdoor video data. Three distinct features are extracted from the contours of detected objects. A unique aggregated feature vector is formed using a fuzzy inference system by aggregating three feature vectors. To minimize computation in recognition using hidden Markov model, the length of final feature vector is reduced using vector quantization.

1.2 Our motivation

As can be seen in state of the art there are many methods capable for human action analysis however most of those papers do not address motion capture data that is generate by modern multimedia motion capture devices that are actual available on the market. Those devices (like for example Microsoft Kinect) is capable to generate low frame rate (about 30 fps) and very noisy stream of tracking data however due to low price that type of data acquisition devices become common components of home multimedia systems. Also those types of devices require different approach in feature selection and signal analysis than RGB or monochrome video images or super-precise wearable motion capture costumes.

The goal of this paper is to propose a reliable way of features selection for motion capture data that will enable further classification of user actions with appropriate pattern recognition algorithm. The main novelty of this paper besides presenting this approach is evaluation of results on large (containing 770 actions) motion-capture dataset of various gym exercises. The common problem with markerless motion capture vision - based systems is that if some parts of body surface are covered by another part it is impossible to perform accurate features measurements. Knowing this we have evaluated two types of features sets: angle-based and coordinates based. We have performed PCA analysis of both features sets to justify our choices. We have also test 24 subsets of proposed features dataset in order to estimate which body joints are vital for correct actions recognition. As a classifier we have evaluated HMM with various number of hidden states and also Gaussian mixture model.

Some of already mentioned methods might be applied to motion capture (MoCap) data acquired by Kinect controller however very often data processing and training procedure

contains many parameters that might be difficult to set-up properly to obtain best recognition results (like for example arbitral number of clusters [18], type of kernel etc.).

The choice of continues HMM was motivated by simplicity of its application to signal with Gaussian densities of observations and its high efficiency. As we will show no others processing of input signal beside features calculation presented in Section 2.1 is required to obtain high recognition rate on our test dataset. For comparison in [5] authors uses actions description called Histogram-of-Oriented-Velocity-Vectors. An algorithm first tracks 3D positions of skeleton joints and extracts the velocity vector associated with each joint. The orientation of the velocity vectors are defined by α and β in a spherical coordinate system. Algorithm then forms a spatial histogram of these vectors by grouping them over orientations. To overcome the problem with actions which do not involve a movement authors uses thresholding schema to distinguish them from non-static actions. That method uses SVM and KNN classifier. Our proposition based on angle-based features similarly to [5] is also scale-invariant, speed-invariant and length-invariant descriptor for human actions represented by 3D skeletons acquired by Kinect however it is far simpler to implement and does not require any arbitrary threshold that might strongly depends to tracking hardware and software.

Also paper [1] presents interesting approach to action recognition task. The depth maps are processed to track human silhouettes by considering temporal continuity constraints of human motion information and compute centroids for each activity based on contour generation. In body joints features, depth silhouettes are computed first through geodesic distance to identify anatomical landmarks which produce joint points information from specific body parts. Then, body joints are processed to produce centroid distance features and key joints distance features. Finally, Self-Organized Map (SOM) is employed to train and recognize different human activities from the features. The mentioned paper proposes original method for computation of body joints features while we are utilizing Kinect SDK method [30]. Both methods however cannot overcome the problem of body occlusion which can only be solved when using multi-view camera system. In our paper we have evaluated proposed HMM approach both on our angle-based features and centroids distance features from paper [1].

2 Methods

The problem we want to solve is classification of gym warm-up and fitness exercises. The exercises were recorded in SKL format and stored in SKL files [27]. SKL recording contains motion capture information: the spatial positions of points on human body (so called joints) that were tracked during recording session. As the tracking device we used Microsoft Kinect (capture frequency was 30 Hz), for raw-data segmentation and tracking we used Microsoft Kinect 1.8 SDK (it generates 20 body joints) and to register SKL recordings Gesture Description Language Studio v1.1 [27]. To solve the above task we utilized angle-based features and classification based on HMM with multivariate Gaussian distribution and also Gaussian mixture model.

2.1 Features selection and preprocessing

The role of feature selection is changing original representation of motion capture data that is three-dimensional coordinates of 20 joints to the other that is more convenient

for further analysis. The original representation has two main drawbacks: it is dependent of relative position of user to camera leans and it is 60-dimensional (3 dimensions of Cartesian frame * 20 joints). The dependence from the camera position virtually prevents method from being usable in real-world scenario. Also the number of dimensions might affect the processing speed of methodology making it not real-time. We have changed position-based representation to angle-based representation (see Fig. 1). In this notation the vertexes of angles are positioned either in some important for movements analysis body joints (like elbows – angle 1 and 2, shoulders – angle 3 and 4, knees – angle 6 and 7) or angles measure position of limbs relatively to each other or relatively to torso. The second type of angles we utilized are angle defined between forearms (angle 5), angle between vector defined by joint between shoulders – joint between hips and thighs (angle 8 and 9) and finally angle between thighs (angle 10).

After the feature selection the data is preprocessed. For each data sample we calculate the Z-score for each feature value:

$$\bar{z} = \frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}} \quad (1)$$

Where \bar{x} are values of single feature in data sample, $\mu_{\bar{x}}$ is a mean of elements in \bar{x} and $\sigma_{\bar{x}}$ is a standard deviation of elements in \bar{x} .

The correctness of angles estimation depends on precision of data tracking. In case of marker less vision-based tracking systems the precision is limited especially when some body parts become invisible from the perspective of camera leans. In this case it is impossible to measure angle value correctly. Due to this fact it is important to investigate which features influence recognition process the most. In the next paragraph we will investigate this influence on different subsets of 10-element features set we defined above. All subsets are presented in Fig. 1.

The second type of features we investigated are the state-of-the-art joints coordinates-based set ([1, 2, 7, 16]). In [1] authors propose so called centroids distance features (CDF) and key joints distance features (KDF). In CDF they compute the distance between body joints and the mean centroid of each activity (totally 15 values). Using KDF they calculate the distance between a set of key joints (i.e., extreme body joints) and the other body joints (i.e., sub-extreme landmarks), there are 54 KDF (together with CDF there are 69 features). We have decided to adopt those features to our needs. We used only 12 CDF (see Fig. 2 - left) because Kinect SDK estimations of hands and feet positions seems to be noisy and due to this unreliable. The centroid point coordinate was replaced by spine joint (with zero index in Fig. 2). We also did not use KDF because 69-dimensional representation of movement seems to be sparse and redundant. We also calculated Z-score (1) of each feature value.

2.2 Signal classification based on HMM with multivariate Gaussian distribution

Continuous HMM denoted as $H = \{S, p, A, B\}$ can be expressed as follows [33]: $S = \{S_1, S_2, \dots, S_q\}$ indicates the states where q is the number of the states. The state of the model at time t can be

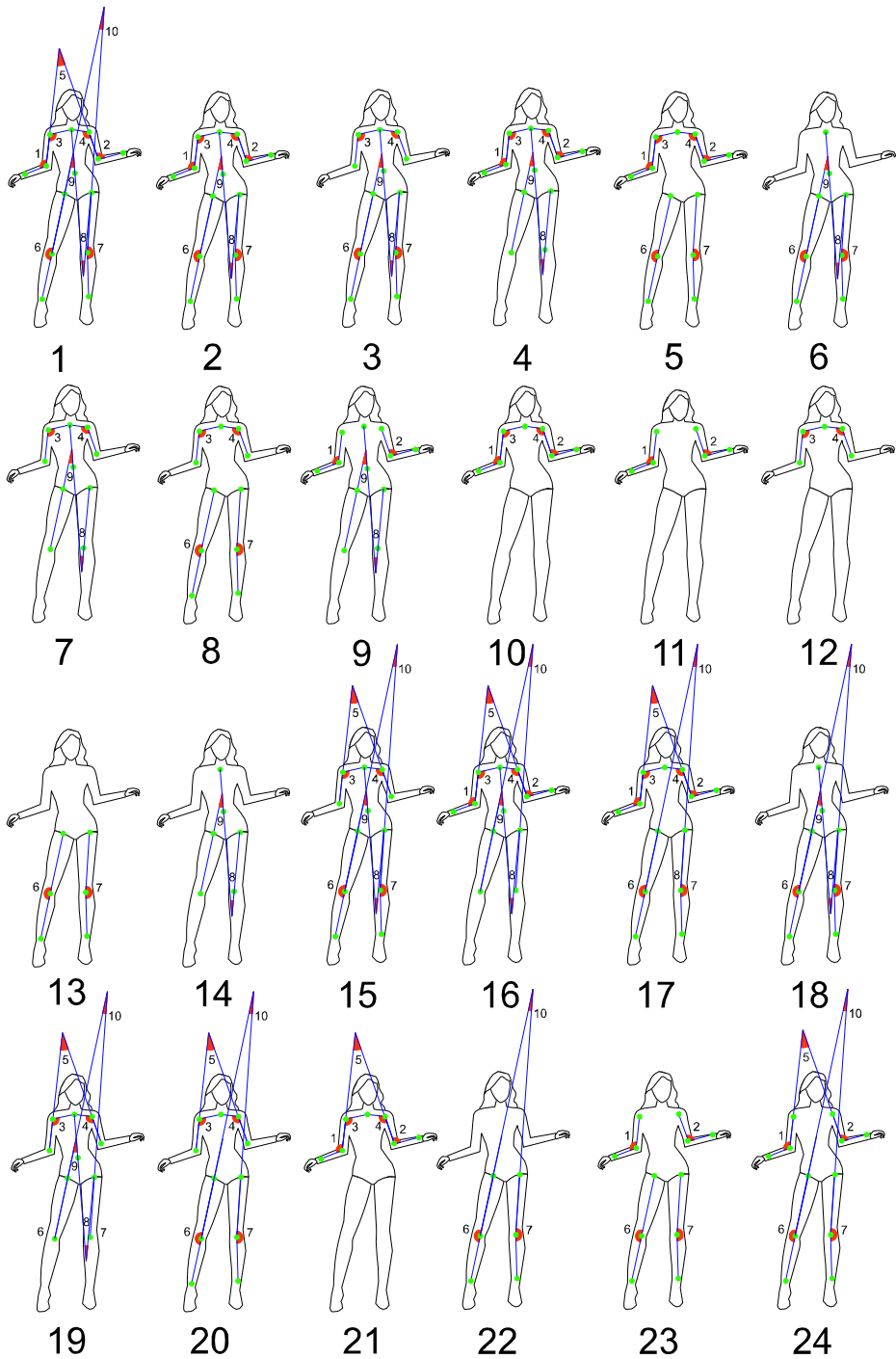
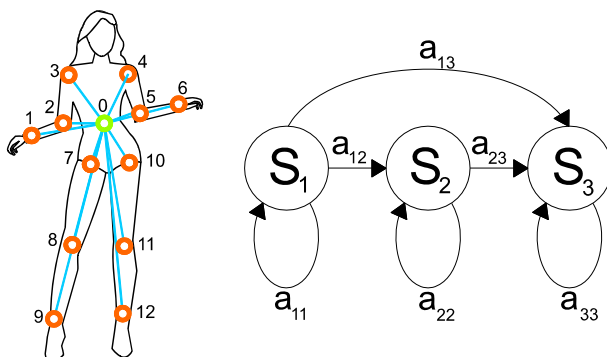


Fig. 1 This figure presents visualization of 24 features sets with tracked points marked as *green circles* and features (angles) marked with *blue lines* and *red semi-circles*. Each feature has its own unique indexing number (from 1 to 10)

Fig. 2 On the left this figure presents visualization of joints coordinates-based feature set we used in our experiment. We compute the distance between spine body joint (index 0) and other body joints (with indexes varying from 1 to 12). On the right this figure presents a three-state left to right HMM



expressed as $\Omega_t \in S, 1 \leq t \leq T$ where T is the length of the observation sequence. The initial probability of the states p can be represented as:

$$p = \{p_j\}, \sum_{j=1}^q p_j = 1 \quad (2)$$

The state transition probability matrix is denoted as A where a_{ij} denotes the probability of a changing state from i to j i.e.,:

$$a_{i,j} = P(\Omega_{t+1} = S_j | \Omega_t = S_i), 1 \leq i, j \leq q \quad (3)$$

$$\sum_{j=1}^q a_{i,j} = 1, 1 \leq i \leq q \quad (4)$$

The observation probability matrix is denoted as B where the probability $b_j(d)$ represents the probability of observing d from a state j that can be expressed as:

$$b_j(d) = P(O_t = d | \Omega_t = S_j), 1 \leq j \leq q \quad (5)$$

In case of forward model (see Fig. 2 - right) the initial probability of the states p are initialized as $\{1, 0, 0\}$ if there are three states and the modeling procedure is to be started from the first state. The transition matrix can be uniformly distributed based on connections between states:

$$A = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 1 \end{bmatrix} \quad (6)$$

In the continuous observation probability density function matrix B , the commonly used distribution to describe the observation densities is the Gaussian one. To represent

the continuous observation probability matrix, the mean and covariance are utilized. Weight coefficients are necessary to use during the mixture of the probability density function (pdf). Thus, the observation probability of O_t at time t from state j can be represented as

$$b_j(O_t) = \sum_{k=1}^M c_{j,k} \cdot b_{j,k}(O_t), 1 \leq j \leq q \quad (7)$$

$$\sum_{k=1}^M c_{j,k} = 1, 1 \leq j \leq q \quad (8)$$

where c represents the weighting coefficients, M the number of mixtures, and O_t the observation feature vector at time t . The Baum-Welch algorithm can be applied to estimate the HMM parameters [3]. In order to find probability of observed sequence a forward algorithm might be used. It is also possible to create a HMM-based classifier that decides to which of n -classes a given signal belongs. In this case the classifier is composed of n HMM (the same as number of classes) and parameters of each HMM are estimated on exemplars of signals from different classes. When unknown signal is being examined it is classified to the class for which HMM corresponding to this class has largest probability to produce the sequence.

3 Experiment and results

The dataset we used in our experiment consists of SKL recordings for 14 participants, 4 women (W1-W4) and 10 men (M1-M10) – W means a woman, M – man, numbers defines id of a participant. The exercises that were performed were: body weight lunge left (bwl), body weight lunge right (bwr), body weight squat (bws), dumbbell bicep curl (dbc), jumping jacks (jj), side lunges left (sll), side lunges right (slr), standing dumbbell upright row (sdur), triceps dumbbell kickback (tdk). In Fig. 3 we present body joints sets (so called skeletons) that represent important phases of each physical activity from our dataset. We have selected those nine actions because they were among exercises that we were capable to register on our multimedia hardware due to trajectories of movements (nearly all tracked joints are visible during whole activity). Also those exercises did not require equipment that might disturb the process of body tracking. Quite often proper equipment handling is crucial part of the exercise's safety and we did not want to incorporate to our dataset actions that might be potentially harmful to experiment participants.

In Table 1 we have presented quantities of actions of a given type that were present in our dataset. As can be seen not every person have performed each action, also numbers of repetitions are not equal. That is because that recordings were made in a certain period of time and not all users were asked to perform all gestures (for example in four recordings bws were skipped). Those lacks were then completed by recordings from other four persons. Each person was asked to perform each exercise how many times he or she is capable to, but not more than ten times each (in order not to get too tired). There were some people who made those exercises more than ten times (for example M1). For the other hand many participants were getting tired more quickly and we decided to

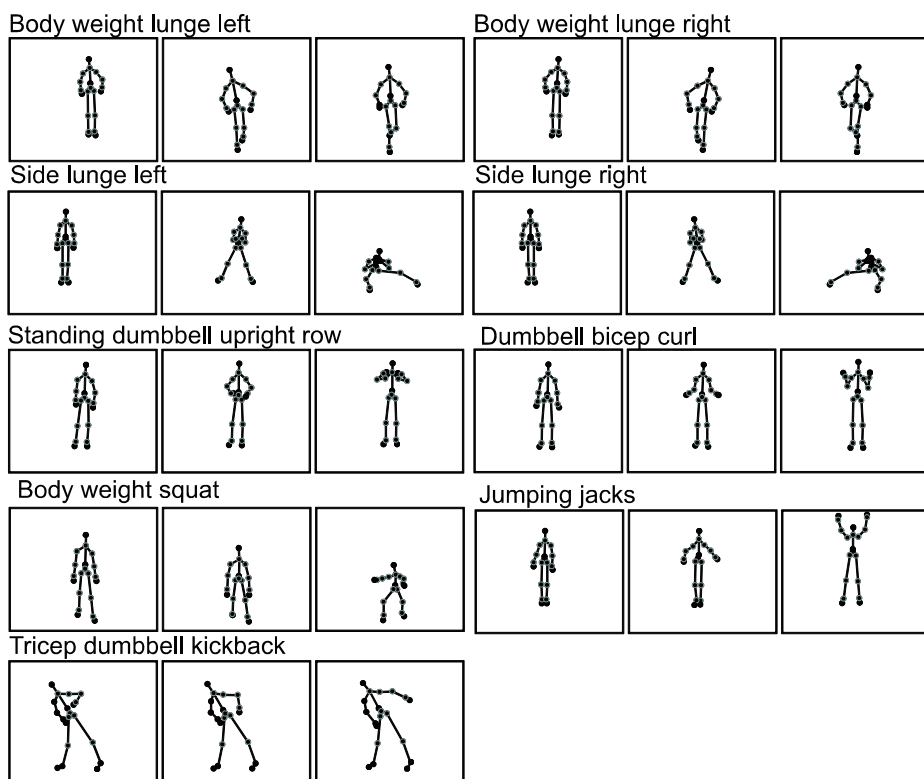


Fig. 3 In this figure we present body joints sets (skeletons) that represent important phases of each physical activity from our dataset

reduce number of repetitions to 5 of each type. For example participant M4 after performing slr was not capable to perform slr correctly. The test set was portioned into 10 groups of recordings (the indexing number of test set is presented in second column of Table 1). The criterion of division was that each set consists of SKL recordings from a single person. The movement samples of this person are not present in other sets. In case of users who did not performed particular action (bws in M3, M4, M5 and M8) those sets were complement with W3, W4, M6 and M9.

The evaluation goes as follow: each HMM classifier composed of q -state HMM was trained to recognize actions presented in Table 1 (each HMM in classifier was trained to recognize single action). We have utilized Baum-Welch learning method. The classifier was trained on nine from ten dataset and then evaluated on remaining one which was test dataset. Each action recording from the test dataset was classified to a class for which HMM corresponding to this class has largest probability to produce the sequence. The HMM classifiers were composed from nine HMM (the same number as classes of actions) that has $q=1, 2, 3$ and then 4 states. A 1-state HMM is basically just a Gaussian mixture model (GMM).

We have implemented our approach in C# utilizing HMM libraries from Accord Framework [31]. Some parts of experiment were evaluated with FactoMineR package [23]. The evaluation of the methodology was performed on dataset presented in Table 1. We have performed K-fold cross validation where $K=10$. After evaluation of each action

Table 1 This table presents the test dataset we used in our experiment. We have used recordings of 14 persons both men and woman (W and M respectively). This partitioning is shown in first column. Those datasets were used to create 10 test datasets for K-fold cross validation. The division on test datasets is shown in second column. Each test dataset besides one (the sixth) contains each type of examined actions. The other columns contain quantities of actions (exercises) of a given type that were performed by a particular person

Dataset	Test set no.	bwl	bwlr	bws	dbc	jj	sll	slr	sdur	tdk
W1	1	10	10	10	10	10	10	10	10	10
W2	2	5	5	5	5	11	5	5	5	6
M1	3	13	10	12	11	9	10	7	12	10
M2	4	10	10	10	10	12	12	9	10	10
M3	5	10	10		10	10	8	10	10	10
W3	5			10						
M4	6	7	5		5	5		5	5	5
W4	6			5						
M5	7	10	10		10	10	5	10	10	10
M6	7			10						
M7	8	10	10	10	10	10	10	10	10	10
M8	9	5	5		5	5	6	5	5	5
M9	9			5						
M10	10	10	10	10	10	10	10	10	10	10
Sum	X	90	85	87	86	92	76	81	87	86

from a test dataset we trained classifier on another nine from ten dataset until all dataset was considered as test dataset. The above procedure was repeated for all features sets presented in Fig. 1.

We have investigated the angle-based and coordinates-based features. We also have taken into account features sets that are derived from both features sets with PCA. We have taken into account datasets with reduced dimensionality from range 3 to 7. We did that because we wanted to investigate the correlation between PC and features to find most relevant for correction recognition subsets. The results of our evaluation are presented in Table 2 and visualized in Fig. 4. The recognition rate (RR) is an averaged value of RR (percentage of correctly recognized actions among action of a given type) among all 10 considered test sets plus/minus standard deviation. The bolded font indicates the settings where averaged RR equals or is above 95 %. We did not present here evaluation of HMMs with number of states greater than 4 because we did not notice any significant increase of RR when number of states is above this value.

In Table 3 and Fig. 5 we present percentage of overall variance for i -th principal component of angle-based and coordinate-based features. This parameter can be calculated from:

$$Var_i = \frac{\lambda_i}{\sum_{j=1}^r \lambda_j} \cdot 100\%$$

Where λ_i is an eigenvalue corresponding to i -th principal component when overall components count is r .

Table 2 Table 4. This table presents results of K-fold cross validation on various forward-only HMM. Each column corresponds to HMM with different number of hidden states. Each row corresponds to different features dataset, Angle means angle-based, Coordinate means coordinates-based. PCA stands for number of principal components of feature set that was used for classification (all means all original features). Values in this table are averaged recognition rate of recognition rate on all action types \pm standard deviation

	States 1 (GMM)	States 2	States 3	States 4
Angle (PCA 3)	60 \pm 25 %	72 \pm 20 %	72 \pm 22 %	74 \pm 20 %
Angle (PCA 4)	76 \pm 17 %	86 \pm 13 %	89 \pm 12 %	88 \pm 11 %
Angle (PCA 5)	80 \pm 15 %	92 \pm 9 %	93 \pm 11 %	94 \pm 9 %
Angle (PCA 6)	83 \pm 11 %	92 \pm 7 %	94 \pm 8 %	95 \pm 9 %
Angle (PCA 7)	90 \pm 9 %	95 \pm 5 %	95 \pm 5 %	96 \pm 6 %
Angle (all 10)	96 \pm 17 %	96 \pm 17 %	96 \pm 15 %	97 \pm 14 %
Coordinate (PCA 3)	53 \pm 27 %	68 \pm 23 %	68 \pm 25 %	68 \pm 28 %
Coordinate (PCA 4)	67 \pm 25 %	74 \pm 23 %	75 \pm 23 %	75 \pm 22 %
Coordinate (PCA 5)	68 \pm 21 %	78 \pm 17 %	79 \pm 17 %	80 \pm 15 %
Coordinate (PCA 6)	71 \pm 18 %	82 \pm 14 %	83 \pm 13 %	84 \pm 13 %
Coordinate (PCA 7)	75 \pm 13 %	83 \pm 12 %	87 \pm 10 %	86 \pm 9 %
Coordinate (all 12)	88 \pm 9 %	94 \pm 6 %	96 \pm 3 %	95 \pm 3 %

In further analysis we do not take into account coordinate-based features because we have obtained better RR results with angle-based set.

In Table 4 and Fig. 6 we present correlations of angle-based features and projection of angle-based features on two-dimensional PC space.

Figure 6 visualize projection of angle-based features on two-dimensional PC space. The PCA supports us with information about correlation between features. The arrows are

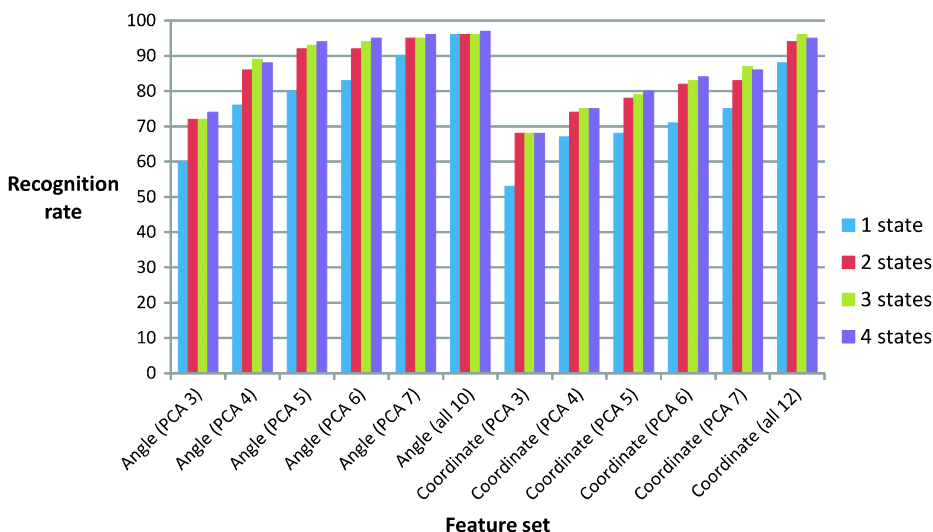


Fig. 4 This figure visualizes data from Table 2 in order to compare recognition rates between classifiers. Color bars corresponds to number of states in HMM classifiers. Vertical axis shows averaged recognition rate of a classifier. Horizontal axis corresponds to different features set

Table 3 This table presents percentage of overall variance for i-th principal component of angle-based and coordinate-based features

	Angle-based features	Coordinate-based features
comp 1	34 %	36 %
comp 2	16 %	21 %
comp 3	13 %	9 %
comp 4	9 %	7 %
comp 5	7 %	6 %
comp 6	6 %	5 %
comp 7	5 %	5 %
comp 8	4 %	4 %
comp 9	3 %	3 %
comp 10	3 %	2 %
comp 11	—	1 %
comp 12	—	1 %

determined by coordinates of the variables in PC frame. Dim 1 contains 34 % of overall variance while Dim 2 16 % (see Tables 3 and 4).

We present detailed evaluation results of 4-states HMM classifier with angle-based features in Table 5. We decided to limit in Tables 2, 3, 4 and 5 presentation of previous GMM / HMM evaluations on proposed combination of features sets to overall recognition rate because our intention was to select the configuration with highest RR. The detailed presentation as in Table 6 will require much more space and it will not be much informative for the reader.

In Table 5 the RR for each classifier with given features set is averaged over all actions and among all 10 considered test sets plus/minus standard deviation. Due to this recognition rate of a classifier is represented by single cell in the Table 5. The bolded font indicates the settings where averaged RR equals or is above 95 %. We did not present here evaluation of HMMs with number of states greater than 4 because we did not notice any significant increase of RR when number of states is above this value (Figs. 7 and 8).

4 Discussion

Our evaluation results in Table 2 and Fig. 4 showed that angle-based features in our dataset guarantee better RR than coordinate-based features. Only in two cases (with

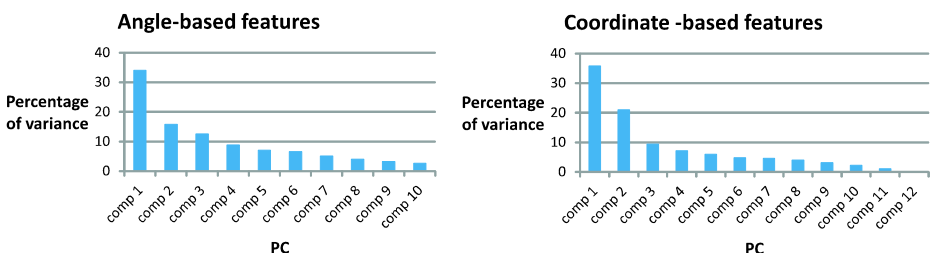
**Fig. 5** This figure visualizes results from Table 3. Vertical axis shows percentage of variance of particular PC while each bar corresponds to different PC

Table 4 This table presents correlations of angle-based features to first PC ordered by corresponding eigenvalue. In brackets in first column there are indexes of angles from Fig. 1

	Dim 1	Dim 2	Dim 3	Dim 4	Dim 5	Dim 6	Dim 7
RightHip (8)	0.76	0.00	−0.13	−0.39	0.11	−0.03	0.35
LeftHip (9)	0.75	0.02	−0.22	0.36	−0.14	0.13	0.37
RightKnee (7)	0.70	0.06	−0.18	−0.50	0.23	−0.12	−0.30
LeftKnee (6)	0.67	0.14	−0.33	0.42	−0.12	0.08	−0.37
RightShoulder (4)	0.47	0.34	0.51	−0.25	−0.47	0.15	−0.01
LeftShoulder (3)	0.41	0.45	0.57	0.19	−0.02	−0.39	−0.06
RightElbow (2)	0.39	−0.55	0.43	0.25	0.32	−0.27	0.06
LeftElbow (1)	0.31	−0.61	0.46	−0.03	0.04	0.49	−0.13
BetweenWrists (5)	0.05	0.73	0.16	0.12	0.52	0.36	0.04
BetweenLegs (10)	−0.83	0.15	0.21	−0.04	−0.02	0.02	0.06

HMM with 3 and 4 hidden states) the recognition rate of coordinate-based features had value greater and equal 95 % while in case of angle-based in 8 cases. While projecting dataset to lower dimension according to PCA angle-based features has always better RR than coordinate-based. However the overall RR in case of using all features without projecting into lower dimension does not differs much (8 % for 1 state for GMM, 2 % for 2 states HMM, 0 % for 3 state HMM and 2 % for 4 states). That means that multivariate continuous hidden Markov model classifier is convenient method of actions recognition. Also less angle-based features are required to obtain better RR. The percentage of overall variance for *i*-th principal component (Table 3 and Fig. 5) is similar in both features set. First four components have over 70 % of overall variance. Table 4 and Fig. 6 shows that BetweenWrists feature (number 5 from Fig. 1) and BetweenLegs (number 10 from Fig. 1) has largest influence on Dim 1 and Dim 2 appropriately. It is also confirmed by evaluation presented in Table 5, where most of features sets (namely 1, 15, 16, 17, 19, 20–24) containing those two features has RR over 90 %.

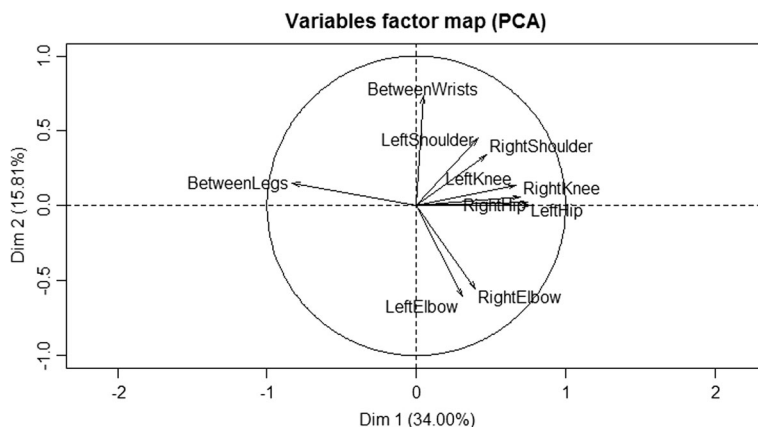
**Fig. 6** This figure visualizes projection of angle-based features on two-dimensional PC space. The data is from Tables 3 and 4

Table 5 This table presents results of K-fold cross validation on various forward-only HMM. Each column corresponds to HMM with different number of hidden states. Each row corresponds to different angle-based features dataset (see Fig. 1). Values in this table are averaged recognition rate of recognition rate on all action types \pm standard deviation

	States 1 (GMM)	States 2	States 3	States 4
1	96 \pm 17 %	96 \pm 17 %	96 \pm 15 %	97 \pm 14 %
2	91 \pm 21 %	95 \pm 15 %	96 \pm 15 %	95 \pm 15 %
3	83 \pm 26 %	91 \pm 16 %	93 \pm 16 %	93 \pm 15 %
4	67 \pm 36 %	87 \pm 23 %	89 \pm 22 %	88 \pm 23 %
5	76 \pm 31 %	88 \pm 25 %	90 \pm 23 %	90 \pm 22 %
6	75 \pm 31 %	88 \pm 20 %	89 \pm 19 %	89 \pm 21 %
7	52 \pm 34 %	77 \pm 26 %	79 \pm 25 %	80 \pm 26 %
8	59 \pm 37 %	81 \pm 24 %	83 \pm 23 %	83 \pm 23 %
9	59 \pm 37 %	81 \pm 27 %	82 \pm 27 %	83 \pm 27 %
10	60 \pm 37 %	73 \pm 31 %	72 \pm 29 %	73 \pm 29 %
11	31 \pm 31 %	51 \pm 31 %	52 \pm 30 %	49 \pm 31 %
12	26 \pm 31 %	45 \pm 29 %	46 \pm 31 %	45 \pm 32 %
13	36 \pm 34 %	68 \pm 29 %	71 \pm 29 %	69 \pm 30 %
14	27 \pm 31 %	56 \pm 35 %	61 \pm 32 %	62 \pm 31 %
15	94 \pm 16 %	96 \pm 15 %	96 \pm 13 %	96 \pm 13 %
16	88 \pm 25 %	93 \pm 21 %	93 \pm 21 %	93 \pm 21 %
17	92 \pm 19 %	95 \pm 18 %	95 \pm 18 %	94 \pm 17 %
18	88 \pm 22 %	92 \pm 17 %	94 \pm 15 %	93 \pm 16 %
19	86 \pm 24 %	90 \pm 23 %	91 \pm 20 %	92 \pm 21 %
20	87 \pm 22 %	93 \pm 19 %	93 \pm 19 %	93 \pm 17 %
21	67 \pm 35 %	77 \pm 29 %	78 \pm 29 %	80 \pm 27 %
22	69 \pm 34 %	80 \pm 27 %	82 \pm 26 %	82 \pm 26 %
23	64 \pm 32 %	83 \pm 26 %	87 \pm 23 %	87 \pm 24 %
24	91 \pm 20 %	95 \pm 16 %	95 \pm 16 %	95 \pm 16 %

As can be seen in Table 5 15 examined settings of classifiers obtained RR greater or equal 95 %. In case we select all ten angle-based features even GMM is capable to correctly identify most (96 \pm 17 %) of the actions. However when number of features is lowered also the RR in GMM drops more quickly than in other considered classifiers. The highest RR 97 \pm 14 % was obtained for 4-states HMM with 10 features (see Tables 5 and 6). The most difficult to classify were actions that required whole body movements and in which some body parts were covered by another ones from the perspective of our tracking hardware. In those cases the data processing software approximated positions of body joints that were not visible for camera's lens. Those approximations however very often did not suit the real joint's positions which affected the further features calculation. In those cases we could not overcome this phenomena by increasing the number of features because they calculation were based on wrongly tracked body joints. Most difficult to classify were mainly sll and slr. They were the only actions where averaged RR in 4-states HMM classifier dropped below 90 %. Sadly, because those errors were caused by limitation of tracking hardware they cannot be easily overcome.

The quite obvious observation is that the more features were used to describe the movement the better recognition results was obtain. Quite surprisingly when only two features were used for recognition (see features set 11, 12 and 13) all classifiers beside

Table 6 This table presents more detailed results of K-fold validation on 4 states forward-only HMM. Each column corresponds to different action. Each row corresponds to different angle-based features dataset (see Fig. 1). Values in this table are 10-fold cross validation results \pm standard deviation

Set	bwll	bwlr	bws	dbc	jf	sdur	sll	slr	tdk	Average
1	100±0 %	98±6 %	98±4 %	100±0 %	100±0 %	99±3 %	87±8 %	89±23 %	100±0 %	97±14 %
2	100±0 %	100±0 %	97±6 %	100±0 %	100±0 %	93±11 %	87±7 %	83±28 %	99±3 %	95±15 %
3	100±0 %	98±4 %	96±7 %	92±7 %	99±3 %	87±14 %	86±8 %	89±18 %	87±19 %	93±15 %
4	94±8 %	98±6 %	84±27 %	97±9 %	93±8 %	88±18 %	73±19 %	72±41 %	97±5 %	88±23 %
5	81±27 %	84±23 %	95±10 %	100±0 %	100±0 %	96±7 %	80±19 %	81±32 %	97±5 %	90±22 %
6	100±0 %	89±22 %	95±9 %	83±13 %	100±0 %	79±28 %	83±13 %	89±20 %	83±25 %	89±21 %
7	80p25 %	86±17 %	81±21 %	80±18 %	94±11 %	87±15 %	58±38 %	71±30 %	84±19 %	80±26 %
8	77±15 %	86±14 %	94±15 %	77±28 %	94±10 %	86±23 %	64±31 %	84±24 %	87±13 %	83±23 %
9	93±7 %	85±23 %	83±28 %	92±11 %	91±12 %	86±20 %	48±36 %	69±32 %	98±4 %	83±27 %
10	51±23 %	84±14 %	79±30 %	96±7 %	73±32 %	80±21 %	46±30 %	71±32 %	79±18 %	73±29 %
11	21±20 %	63±30 %	51±22 %	80±21 %	63±22 %	35±26 %	22±20 %	45±25 %	59±30 %	49±31 %
12	14±21 %	40±26 %	50±29 %	41±30 %	76±29 %	49±29 %	27±24 %	63±26 %	51±24 %	45±32 %
13	85±19 %	70±22 %	93±12 %	40±24 %	77±23 %	81±19 %	54±26 %	68±33 %	51±34 %	69±30 %
14	77±27 %	69±39 %	83±23 %	36±19 %	55±16 %	74±22 %	28±26 %	58±27 %	76±17 %	62±31 %
15	100±0 %	98±6 %	98±4 %	100±0 %	99±3 %	96±6 %	88±4 %	87±20 %	99±3 %	96±13 %
16	100±0 %	98±6 %	100±0 %	100±0 %	98±4 %	97±6 %	73±27 %	74±39 %	99±3 %	93±21 %
17	100±0 %	99±3 %	100±0 %	100±0 %	99±3 %	97±6 %	84±11 %	77±34 %	94±12 %	94±17 %
18	100±0 %	94±10 %	98±4 %	90±9 %	99±3 %	89±15 %	86±9 %	88±18 %	91±18 %	93±16 %
19	98±6 %	98±6 %	98±6 %	95±7 %	97±6 %	96±6 %	72±31 %	70±37 %	99±3 %	92±21 %
20	97±6 %	97±6 %	97±6 %	94±10 %	99±3 %	100±0 %	79±17 %	80±26 %	91±19 %	93±17 %
21	73±28 %	69±17 %	97±6 %	99±3 %	86±20 %	96±7 %	55±28 %	71±32 %	78±31 %	80±27 %
22	94±8 %	83±25 %	93±13 %	62±24 %	99±3 %	84±16 %	77±25 %	81±24 %	63±33 %	82±26 %
23	79±31 %	90±30 %	92±12 %	95±12 %	94±8 %	92±9 %	65±30 %	78±24 %	95±6 %	87±24 %
24	100±0 %	99±3 %	97±6 %	100±0 %	100±0 %	96±7 %	81±16 %	83±28 %	99±3 %	95±16 %

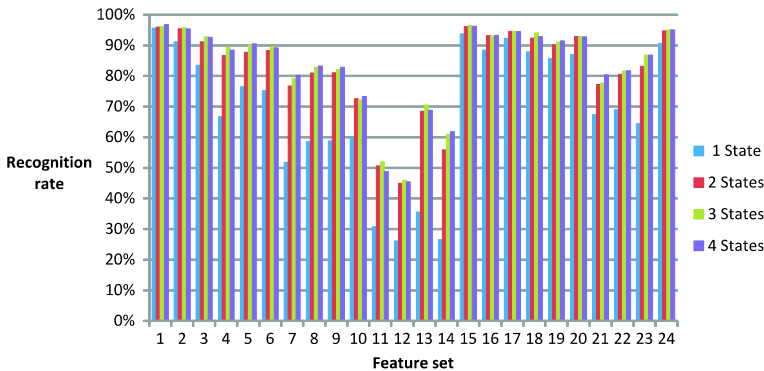


Fig. 7 This figure visualizes data from Table 5 in order to compare recognition rates between classifiers. *Color bars* corresponds to number of states in HMM classifiers. *Vertical axis* shows averaged recognition rate of a classifier. *Horizontal axis* corresponds to different angle-based features set (see Fig. 1)

GMM were capable to recognize nearly or over 50 % actions. The more important for correct recognition seemed to be knee joints that those in elbows. The bigger impact on averaged RR that leg-based features had over hand-based could also be observed in features sets 6 and 7.

The very important observation is that it is possible to obtain average RR for HMM with more than 1 state at level 95 % without using the feature that are based on joints positioned in torso of the observed user (see features set 24). Also in features set 24 RR does not differ much (1 to 2 %) from the best results from features set 1. This is very important information because during physical activities torso is very often covered from the camera's perspective by limbs. We can utilize this knowledge when designing data acquisition set up for action recognition and track only body joints positioned on limbs (hands and legs).

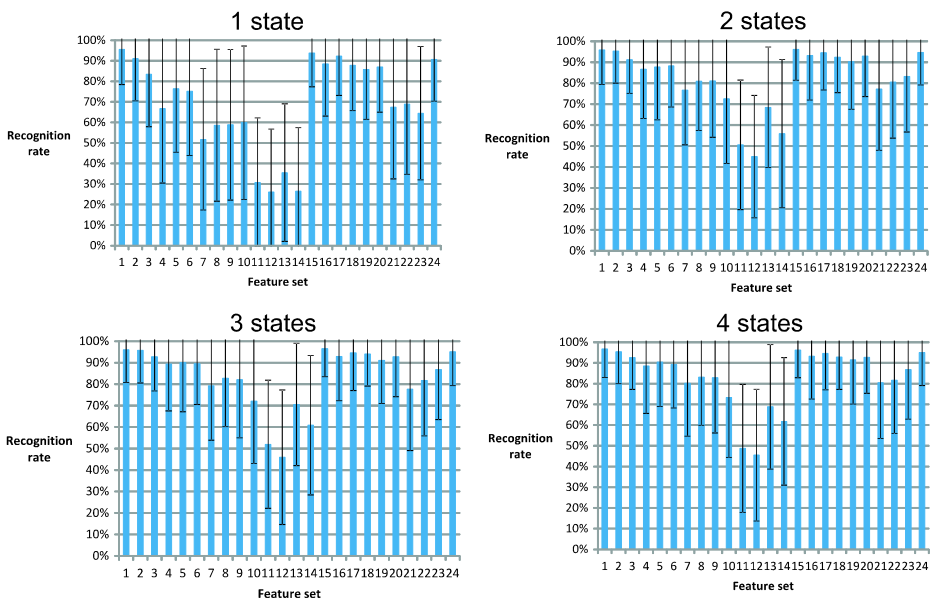


Fig. 8 This figure visualize results from Table 5 in order to show averaged recognition rates (*blue bars*) and standard deviations (*black bars*). Each plot corresponds to different forward-only HMM classifier. *Vertical axis* shows averaged recognition rate of a classifier. *Horizontal axis* corresponds to different angle-based features set (see Fig. 1)

As can be seen highest RR was achieved in nearly all features set for HMM with 4 states however difference between 4, 3 and 2 states HMM is not huge – about 1 or 2 %. Also The standard deviation of those results are nearly identical – they did not differ more than 4 % (in case of features set 14) but more often differ from 1 to 2 %. That means that classifiers react in very similar way on lowering/altering observed features. The implementation of forward algorithm in case of all considered HMM classifiers enables real time performance (with frequency above 30 Hz) on our dataset. Knowing that we can safely assume that forward-only 4 states HMM – based classifier is suitable to solve our classification problem.

5 Conclusion

In this paper we describe how to create a HMM-based classifier that is capable of real-time gym actions recognition. What is more we proved that this classifier can operate on low-quality motion capture data generated by multimedia hardware. The features sets we used are invariant to relative position of user to camera leans. With the angle-based features we proposed it is even possible to obtain high RR on simple GMM classifier. The experimental results we have observed on angle-based features subsets were confirmed by PCA analysis. The features that have more variance (that influence the “most relevant” PC the most) are also present in subsets with highest RR.

The methodology and results presented in this paper are significant both from practical and theoretical point of view. Practically the implementation of our methodology can be directly applied to multimedia computer coaching application. The goal of such program is to oversee the gym warming up and training by checking if a person follows a workout routine, to show its progress and to motivate for further efforts. However we have to remember that it is very difficult or nearly impossible to apply proposed methodology to evaluate the quality (similarity to “ideal pattern”) of gym exercises when we are using only multimedia hardware. We believe that the qualitative evaluation will be possible after applying professional motion capture costume especially that professional hardware also produces very similar set of body joints as Kinect does.

The results of our work also indicate which angle-based features sets are most important for further correct recognition of body actions. That means which parts of human body should be captured in order to optimize the RR of the observed actions. Those findings might be then applied for example during designing of outdoor real-time system which detects potentially hazardous situation that might lead to violence. This type of solution is extension to street monitoring systems. With our results there is a possibility to optimize the position of cameras to enable monitoring of those parts of human figure that are most important for further behavior classification. Also knowing our findings it will be possible to limit the number of sensor on motion capture costume if the intention is to use this type of hardware to action recognition and quality evaluation. The optimization of position and limitation of sensors significantly lowers the cost of hardware because motion capturing sensors are relatively expensive part of equipment.

The goal of our further researches will be to utilize the results presented in this paper to two already mentioned applications that is outdoor real-time hazardous situation detection and high-quality body actions analysis (especially in sport). Basing on latest achievement in web-based visual data computation and high-quality visualization [28, 32] we believe that there are no technological and methodological obstacles for solving our scientific goals and to introduce the human action recognition even to low-power mobile devices.

Acknowledgments We kindly thank company NatuMed Sp. z o.o (Targowa 17a, 42-244 Wancerzow, Poland) for supplying us with SKL dataset that together with our own SKL recordings was used as training and validation dataset in this research.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- (2014) Depth map-based human activity tracking and recognition using body joints features and self-organized map, computing, communication and networking technologies (ICCCNT). Int Conf IEEE 1–6, doi:[10.1109/ICCCNT.2014.6963013](https://doi.org/10.1109/ICCCNT.2014.6963013)
- Arici T, Celebi S, Aydin AS, Temiz TT (2014) Robust gesture recognition using feature pre-processing and weighted dynamic time warping. *Multimedia Tools Appl* 72(3):3045–3062
- Baum E, Petrie T, Soules G, Weiss N (1970) A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains. *Ann Math Stat* 41:164–171
- Beh J, Han D, Ko H (2014) Rule-based trajectory segmentation for modeling hand motion trajectory. *Pattern Recog* 47(4):1586–1601
- Boubou S, Suzuki E (2015) Classifying actions based on histogram of oriented velocity vectors. *J Intell Inf Syst* 44(1):49–65
- Chen F-S, Fu C-M, Huang C-L (2003) Hand gesture recognition using a real-time tracking method and hidden Markov models. *Image Vision Comput* 21:745–758
- Chen X, Koskela M (2015) Skeleton-based action recognition with extreme learning machines. *Neurocomputing* 149(Part A):387–396
- Game N, Kuang YC, Akmeliawati R, Demidenko S (2011) Gaussian process dynamical models for hand gesture interpretation in sign language. *Pattern Recog Lett* 32(15):2009–2014
- Głowacz A, Głowacz W (2008) Dc machine diagnostics based on sound recognition with application of FFT and fuzzy logic. *Przegl Electrotech* 84(12):43–46
- Głowacz A, Głowacz W (2008) Sound recognition of dc machine with-application of FFT and backpropagation neural network. Conference: 13th International Power Electronics and Motion Control Conference Location: Poznan, POLAND Date: SEP 01-03, 2008. *Przegl Electrotech* 84(9):159–162
- Hachaj T, Ogiela MR (2014) Full-body gestures and movements recognition: user descriptive and unsupervised learning approaches in GDL classifier. Conference: Conference on Applications of Digital Image Processing XXXVII Location: San Diego, CA Date: AUG 18-21, 2014, SPIE, APPLICATIONS OF DIGITAL IMAGE PROCESSING XXXVII Book Series. Proc SPIE 9217:921704, Published: 2014
- Hachaj T, Ogiela MR (2014) Rule-based approach to recognizing human body poses and gestures in real time. *Multimedia Systems* 20(1):81–99
- Huang C-L, Jeng S-H (2001) A model-based hand gesture recognition system. *Mach Vis Appl* 12(5):243–258
- Huang C-L, Wu M-S, Jeng S-H (2000) Gesture recognition using the multi-PDM method and hidden Markov model. *Image Vis Comput* 18(11):865–879
- Ibañez R, Soria Á, Teyseyre A, Campo M (2014) Easy gesture recognition for Kinect. *Adv Eng Softw* 76:171–180
- Ibañez R, Soria Á, Teyseyre A, Campo M (2014) Easy gesture recognition for Kinect. *Adv Eng Softw* 76:171–180
- Kanga H, Leea CW, Jung K (2004) Recognition-based gesture spotting in video games. *Pattern Recog Lett* 25(15):1701–1714
- Khan Zafar A, Sohn W (2013) A hierarchical abnormal human activity recognition system based on R-transform and kernel discriminant analysis for elderly health care. *Computing* 95(2):109–127
- Kim I-C, Chien S-I (2001) Analysis of 3D hand trajectory gestures using stroke-based composite hidden markov models. *Appl Intell* 15(2):131–143
- Kim JH, Kim KK, Suen CY (2000) An HMM-MLP hybrid model for cursive script recognition. *Pattern Anal Appl* 3(4):314–324
- Kim D, Song J, Kim D (2007) Simultaneous gesture segmentation and recognition based on forward spotting accumulative HMMs. *Pattern Recog* 40(11):3012–3026
- Ko AH-R, Sabourin R, Alceu de Souza B Jr (2009) Ensemble of HMM classifiers based on the clustering validity index for a handwritten numeral recognizer. *Pattern Anal Applic* 12:21–35. doi:[10.1007/s10044-007-0094-6](https://doi.org/10.1007/s10044-007-0094-6)
- Lê S, Josse J, Husson F (2008) FactoMineR: an R package for multivariate analysis. *J Stat Softw* 25(1):1–18

24. Lee H, Lim SY, Lee I, Cha J, Cho D-C, Cho S (2013) Multi-modal user interaction method based on gaze tracking and gesture recognition. *Signal Process Image Commun* 28(2):114–126
25. Lee T, Zheng F, Wu W, Chen D (2000) The Hidden Markov model of co-articulation and its application to the continuous speech recognition. *J Electron (China)* 17(3):242–247
26. Mahapatra A, Mishra TK, Sa PK, Majhi B (2014) Human recognition system for outdoor videos using Hidden Markov model. *AEU Int J Electron Commun* 68(3):227–236
27. Official website of GDL technology <http://cci.up.krakow.pl/gdl/>
28. Piorkowski A (2014) Parallel processing in web-based interactive echocardiography simulators. *Comput Inf* 33(3):707–720
29. Qian X, Wang H, Liu G, Hou X (2012) HMM based soccer video event detection using enhanced mid-level semantic. *Multimedia Tools Appl* 60(1):233–255
30. Shotton J, Girshick R, Fitzgibbon A, Sharp T, Cook M, Finocchio M, Moore R, Kohli P, Criminisi A, Kipman A, Blake A (2013) Efficient human pose estimation from single depth images, decision forests for computer vision and medical image analysis. *Adv Comput Vis Pattern Recog* 175–192, doi:[10.1007/978-1-4471-4929-3_13](https://doi.org/10.1007/978-1-4471-4929-3_13)
31. Souza CR (2012) The Accord.NET Framework <http://accord.googlecode.com>
32. Szostek K, Piorkowski A (2010) OpenGL in multi-user web-based applications, Edited by: Sobh, T; Elleithy, K, Conference: International Conference on Systems, Computing Sciences and Software Engineering (SCSS) is part of the International Joint Conferences on Computer, Information, and Systems Sciences, and Engineering (CISSE) Location: Bridgeport, CT Date: DEC 04-12, 2009, Sponsor(s): Inst Elect & Elect Engineers (IEEE); Univ Bridgeport, Innovations in computing sciences and software engineering, Pages: 379-383, Published: 2010
33. Uddin Z, Kim T-S (2001) Continuous Hidden Markov models for depth map-based human activity recognition, computer and information science, numerical analysis and scientific computing, Hidden Markov models, theory and applications, book edited by Przemyslaw Dymarski, ISBN 978-953-307-208-1, Published: April 19, 2011 under CC BY-NC-SA 3.0 license. © The Author(s)
34. Uddin MZ, Lee JJ, Kim T-S (2010) Independent shape component-based human activity recognition via Hidden Markov model. *Appl Intell* 33:193–206. doi:[10.1007/s10489-008-0159-2](https://doi.org/10.1007/s10489-008-0159-2)
35. Yoon H-S, Soh J, Younglae BJ, Yang HS (2001) Hand gesture recognition using combined features of location, angle and velocity. *Pattern Recogn* 34:1491–1501
36. Zarrouk E, Ayed YB, Gargouri F (2014) Hybrid continuous speech recognition systems by HMM, MLP and SVM: a comparative study. *Int J Speech Technol* 17(3):223–233
37. Zhou Y, Cheng Z, Jing L, Wang J, Huang T (2014) Pre-classification based hidden Markov model for quick and accurate gesture recognition using a finger-worn device. *Appl Intell* 40(4):613–622



Dr. Tomasz Hachaj Works in Chair of Computer Science and Computational Methods at the Pedagogical University in Krakow. In 2010 for his doctoral thesis on pattern recognition techniques in semantic interpretation of selected lesions based on analysis of brain tissues blood perfusion visualization he was awarded the title of Doctor of Computer Science at the Faculty of Electrical, Automatic Control, Computer Science and Electronic Engineering of the AGH University of Science and Technology. Author of more than 40 scientific international publications on pattern recognition, image processing and computer graphic.



Prof. Marek R. Ogiela Professor of Computer Science, cognitive scientist and cryptographer. Head of Cryptography and Cognitive Informatics Laboratory in Krakow, Poland. He is a member of numerous world scientific associations i.e., SPIE Senior Member, IEEE Senior Member etc. as well as member of the Interdisciplinary Scientific Committee of the Polish Academy of Arts and Sciences. Author of more than 300 scientific international publications on pattern recognition and image understanding, artificial intelligence, cybernetics, cryptography and information theory.